Information Fusion in the Sensorweb

Prasanta Bose, Mark Woodward, Neal Hurlburt, & Sam Freeland

Advanced Technology Center
Lockheed Martin,
Palo Alto, CA 940304-1911

Abstract

Future NASA missions involving constellation of spacecrafts (e.g. Leonardo [9], MAGCON [7]), offers a unique opportunity for better coordinated scientific observations and deeper understanding of the earth, sun, and their connections as a set of interacting systems. Such missions also present several unique challenges to make the data collection and dissemination system to work. These include the need to task multiple sensors and systems and the need to integrate, fuse and present this data to researchers in a seamless, timely and efficient manner. NASA has advanced the vision of the Sensorweb to develop network-centric sensing, fusion and dissemination capabilities required to meet the challenges. This paper develops the concept of the Sensorweb Virtual Machine (SVM), SVM enabled sensorweb nodes (S-Nodes), and agent-based mechanisms for realizing the SVM capabilities. The paper briefly describes coordination of S-Nodes in the context of a Living with Star application.

Keywords: sensorweb, coordination, semantic mark-up languages, architecture, agents

1 Introduction

Future NASA missions involving constellation of spacecrafts (e.g. Leonardo, MAGCON), offer a unique opportunity for better coordinated scientific observations and deeper understanding of the earth, sun, and their connections as set of interacting systems. Such missions also present several unique challenges to make the data collection and dissemination system to work. These include the need to task multiple sensors and systems and the need to integrate, fuse and present this data to researchers in a seamless, timely and efficient manner. Consider the following scenarios involving cooperating distributed nodes/scientists in the context of the Living with Star [8] mission:

- A magnetospheric physicist wishes to compare proton fluxes measured in the magnetosphere to particular solar events observed by the Solar Sentinels and Solar Dynamics Observatory (SDO). This could quickly lead to a complex, interrelated set of searches involving data sets located throughout the country.
- A solar physicist might wish to download a movie sequence from SDO. Since the current design calls for a data rate of 1 terabyte/day even a modest request could overwhelm the user and place severe demands on the server or network.
- A forecaster may wish to obtain a regular set of observations from multiple missions, and to be notified when specific criteria based upon multiple instruments are met.

1.1 Requirements

Developing an infrastructure that support the above scenarios must meet the following major requirements:

- Task-driven automated or semi-automated synthesis of analysis and presentation mechanism. We need the ability to assemble analysis services or routines based on high-level task needs. The services may be provided by a distributed set of analysis tools that are developed and maintained separately.
- Assembling distributed image resources: We need tools for assembly of ***relevant*** distributed data resources for analysis. The data may re

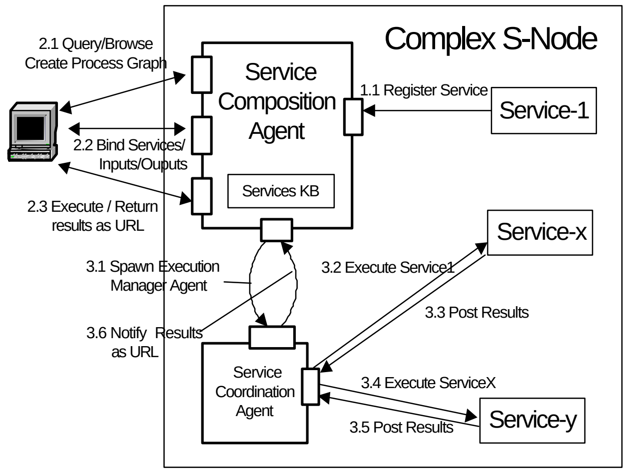


Figure 1: S-Node Elements and their interactions.

- side at different sites in different formats and coordinate systems, and require various levels of calibration, registration, etc.
- Coordinated analysis spanning multiple science domains. We need tools that can react to changes in data or user requirements.
- Coordinated and concurrent execution of services: Given process models for realizing a user task, coordination capabilities are required to execute the process concurrently with other processes. The coordination must consider resource management issues as well as security issues.
- Scaleable architecture enabling hierarchical information fusion.

1.2 Key Ideas

This paper describes our approach to developing a sensorweb infrastructure that meets the above requirements. The two key ideas in our approach are: a) Sensorweb Virtual Machine (SVM) - that provide agent-based mechanisms to mediate and coordinate interactions between the data sources, sensors, processing services and the end users (scientists, analysts, etc.). The SVM enables search, access and analysis of data intelligently across large, distributed data sources as well as enable exploitation of analysis services developed by individual PI's relevant to his/her domain. The SVM enables a node's data, sensor, resources, and services to get exploited by another node and vice-versa exploit other nodes for its tasks. b) Semantic markup of resources, processing services, and data using DAML – the semantic markup enables agents to dynamically discover, compose, invoke, access and use services, data and resources to meet high-level task descriptions of the end user.

In the following sections, we briefly articulate the SVM architecture and describe some of its key elements.

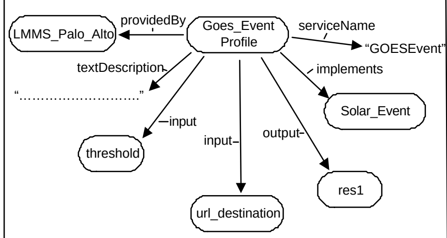
2 Approach

2.1 SVM Architecture,

The key elements of the architecture are:

- **SVM Enabled Sensorweb Nodes (S-**Nodes). S-Nodes provide an abstraction for distributed intelligent fusion, task-driven coordination and resource management. A key element of the S-Node is a set of software agents that forms the sensorweb virtual machine. The agents realize protocols and provides capabilities for 1) Registering resources and services, 2) Supporting interactive task generation, query and obtaining fusion results. 3) Workflow (distributed or centralized) based coordinated execution of process models of tasks. The agents exploit knowledge obtained from semantic descriptions of the services and resources to realize the SVM capabilities. The analysis can be performed based on external requests delegated to it or based on long-standing, local directives.
- **Peer-to-peer or hierarchical coordination architecture.** S-Nodes can enter into peer-to-peer interactions for tasking, distributed and reactive analysis. S-Nodes can also be organized into hierarchical structures,

whereby a higher-level S-Node delegates tasks to a lower-level S-Node. Such interaction relationship is essential for development of scalable sensorwebs.




```
<rdf:RDF
  <lm_concepts:Profile rdf:ID="Goes_Event">
    <lm_concepts:implements rdf:resource="&lm_concepts;#Solar_Event"/>
    <profile:serviceName>GOESEvent</profile:serviceName>
    <profile:textDescription>
      ................
    </profile:textDescription>

    <profile:providedBy>
      <profile:ServiceProvider rdf:ID="LMMS_Palo_Alto">
        <profile:name>Lockheed Martin</profile:name>
      </profile:ServiceProvider>
    </profile:providedBy>

    <profile:input>
      <lm_concepts:ResultURL>
                        <profile:parameterName="url_destination"/>
                        <lm_concepts:textDescription>
                          Post result URL to:
                        </lm_concepts:textDescription>
      </lm_concepts:ResultURL>
    </profile:input>

    <profile:input>
      <lm_concepts:Parameter>
        <profile:parameterName="threshold"/>
      </lm_concepts:Parameter>
    </profile:input>

    <profile:output>
      <lm_conceps:Time>
        <lm_concepts:name>res1</lm_concepts:name>
      </lm_concepts:Time>
    </profile:output>

  </lm_concepts:Profile>
</rdf:RDF>
```

Figure 2: a) A partial Semantic Description of a GOES Service, and b) DAML Description.

2.2 Semantic Description of Services

A major problem is dynamically discovering, configuring and coordinating the resources, data, processing and other services such as sensing and actuation processes (where the node is actually a space-craft or ground-node representative of the s/c) encapsulated by a S-node. One approach, as taken in the Computing Grid [5, 6] is to define a set of standard interfaces that is provided by the services. We take a declarative approach and exploit the semantic markup language [2], DAML (DARPA Agent Markup Language), for representing capabilities provided by a service, process model of the service and computational mechanism model for usage of the service (e.g. whether the service supports messages, events or calls). The representation of such information is based on exploitation of relevant ontologies (or conceptual schemas). For example, figure 2a is a conceptual model of a service provided by an S-Node. Note that the displayed attributes are a subset of the actual attributes. The DAML representation of the conceptual service description in figure 2a is given in figure 2b. In the simplest form of DAML (as well as in Resource Description Format or RDF [5]), every relationship is broken up into a triplet (Object, Attribute, Value). For example, the concept that the service name of the Goes_Event Service is "GOESEvent" is depicted in DAML as,

```
<lm_concepts:Profile rdf:ID="Goes_Event">
        <pr file:serviceName>GOESEvent</profile:serviceName>
</lm_concepts:Profile>
```

where the Object is "Profile" the attribute is "serviceName" and the value is "GOESEvent". A more detailed description (not shown here) would further classify the tags used (e.g "lm_concepts") as subtypes of other concepts in a domain specific ontology (e.g image processing) or in some domain independent ontology (such as propositional logic, state machines, etc.).

3 Initial Implementation

We have implemented a core agent-based SVM to develop an initial system for composition and execution of services and tools for analysis of solar data. The system exploits the extensive set of software tools, which are already deployed. The first key component is Solar Soft, the large collection of programs and techniques developed by the solar community. This software underlies many of the data processing and analysis systems in solar physics and is being developed by Lockheed Martin Solar and Astrophysics Laboratory under funding from NASA.

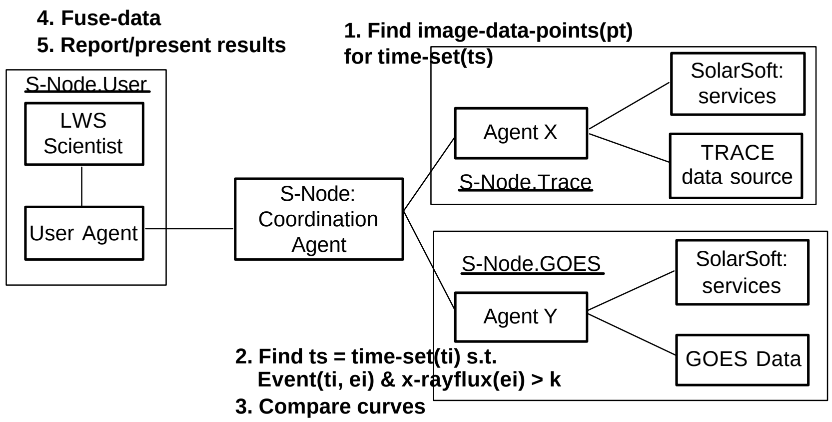


Figure 3: Cooperation of agents realizing the SVM of the Nodes in a LWS application.

3.1 A Scenario

An example behavior of the working system is shown in Figure 3, where the data required for a specific analysis task is distributed.

In this example, the LWS scientists are seeking to identify sources of X-ray flares using the combined datasets. They work with the User-Agent representative to specify their task of identifying X-ray flares by combining TRACE and GOES datasets. The user-agent coordinates with other sensor-agents to do the analysis. This involves:

1. The Agent-x for the GOES S-Node execute appropriate SolarSoft routines to search the GOES time series data and find all events that exceeded a particular X-ray flux.
2. The Agent-y for TRACE S-Node uses the times of excess flux to search the TRACE archive. First they test to see if appropriate data is available. Then they retrieve it, resample it (to reduce transfer and compute times) and generate light curves for all pixels.
3. The Agent-x searches for matches by executing cross-correlation SolarSoft routines.
4. The Agent for user S-Node takes the results from the two agents, and generates summary reports for the scientist.
5. A coordination S-Node mediates the interactions between the different S-Nodes for security.

4 Related and Future Work

The research presented in this paper builds on work in the area of semantic web, software agents and e-science (Computational Grid). We extend and apply the semantic markup language work ([1,2] to consider information services (as opposed to just processing services) that enables intelligent data integration. The work on developing agent-based mechanisms for co-ordination is similar in spirit to the OAA work [3,4] of object integration. The key difference is in terms of the underlying coordination model – we exploit a workflow-based model.

Another key related research is the Computational grid [5, 6]. The Grid is based on defining a standard set of API for each service and developing tools to provide management of service instances based on exploiting the API. For example, service components support API on discovery, authentication, etc. We take a declarative approach that is necessary for task driven search, composition and coordination of services. Our approach can be extended to exploit the Grid capabilities in a complimentary manner. We are currently extending our concepts and SVM agents to exploit the grid service.

5 References

[1] A. Ankolenkar, et. al, "DAML-S: A Semantic Markup Language For Web Services". *Proceedings of The International Semantic Web Working Symposium (SWWS)*, Stanford, CA, July 2001.

[2] T. Berners-Lee, J.Hendler, and O. Lassila. "The semantic web". Scientific American, 284(5):34-43, 2001.

[3] D. Martin, et. al. (1997). "*Information brokering in an agent architecture*". Proceedings of the Second International Conference on the Practical Application of Intelligent Agents and Multi-Agent Technology, (London), The Practical Application Company Ltd., April 1997.
[4] D. L. Martin, et. al., "The Open Agent Architecture: A framework for building distributed software systems," Applied Artificial Intelligence: An International Journal. Volume 13, Number 1-2, January-March 1999. pp 91-128.
[5] Jon Postel and Joe Touch. Network infrastructure. In Ian Foster and Carl Kesselman, editors, The Grid:
Blueprint for a New Computing Infrastructure, chapter 21, pages 552–562. Morgan Kaufmann Publishers, Inc., 1999.
[6] David De Roure, Nicholas Jennings, Nigel Shadbolt and Mark Baker, "Research Agenda for the Semantic Grid: A Future e-Science Infrastructure", Technical Paper, 19th February 2002, http://www.semanticgrid.org/html/semgrid.html
[7] Magnetospheric Constellation STDT Report, NASA, April 2001,
http://stp.gsfc.nasa.gov/missions/mc/mc_STDT_report.htm
[8] G. Withbroe, NASA HQ. "Living with a Star: The Sun-Earth-Connection".
http://lws.gsfc.nasa.gov/docs/old/lws_public.pdf
[9] M. Schoeberl et.al.,"Intelligent Distributed Spacecraft Infrastructure", http://nmp.jpl.nasa.gov/workshop-eo4/proceedings/ESE_Wkshop_ISDCI.pdf